RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/762,769A
Source: IFW16
Date Processed by STIC: 7/12/06

ENTERED

IFW16

RAW SEQUENCE LISTING DATE: 07/12/2006
PATENT APPLICATION: **US/10/762,769A** TIME: 09:52:58

Input Set : **F:\SEQLIST.TXT**
Output Set: **N:\CRF4\07122006\J762769A.raw**

<110> APPLICANT: Melis, Anastasios
Wintz, Hsu-Ching Chen
<120> TITLE OF INVENTION: MODULATION OF SULFATE PERMEASE FOR PHOTOSYNTHETIC HYDROGEN PRODUCTION
<130> FILE REFERENCE: BERK-016CIP
<140> CURRENT APPLICATION NUMBER: 10/762,769A
<141> CURRENT FILING DATE: 2004-01-21
<150> PRIOR APPLICATION NUMBER: 60/354,760
<151> PRIOR FILING DATE: 2002-02-04
<150> PRIOR APPLICATION NUMBER: 60/377,902
<151> PRIOR FILING DATE: 2002-05-02
<150> PRIOR APPLICATION NUMBER: 10/350,298
<151> PRIOR FILING DATE: 2003-01-22
<160> NUMBER OF SEQ ID NOS: 16
<170> SOFTWARE: FastSEQ for Windows Version 4.0
<210> SEQ ID NO: 1
<211> LENGTH: 411
<212> TYPE: PRT
<213> ORGANISM: Chlamydomonas reinhardtii
<400> SEQUENCE: 1

```
Met Glu Arg Val Cys Ser His Gln Leu Ala Ser Ser Arg Gly Arg Pro
 1               5                  10                  15
Cys Ile Ala Gly Val Gln Arg Ser Pro Ile Arg Leu Gly Thr Ser Ser
            20                  25                  30
Val Ala His Val Gln Val Ser Pro Ala Gly Leu Gly Arg Tyr Gln Arg
        35                  40                  45
Gln Arg Leu Gln Val Val Ala Ser Ala Ala Ala Ala Ala Ala Phe Asp
    50                  55                  60
Pro Pro Gly Gly Val Ser Ala Gly Phe Ser Gln Pro Gln Gln Gln Leu
65                  70                  75                  80
Pro Gln Gln His Pro Arg Gln Pro Gln Ala Val Ala Glu Val Ala Val
                85                  90                  95
Ala Glu Ser Val Ser Ala Pro Ala Ser Ala Ala Pro Ser Asn Asp Gly
            100                 105                 110
Ser Pro Thr Ala Ser Met Asp Gly Gly Pro Ser Ser Gly Leu Ser Ala
        115                 120                 125
Val Pro Ala Ala Ala Thr Ala Thr Asp Leu Phe Ser Ala Ala Ala Arg
    130                 135                 140
Leu Arg Leu Pro Asn Leu Ser Pro Ile Ile Thr Trp Thr Phe Met Leu
145                 150                 155                 160
Ser Tyr Met Ala Phe Met Leu Ile Met Pro Ile Thr Ala Leu Leu Gln
                165                 170                 175
Lys Ala Ser Leu Val Pro Leu Asn Val Phe Ile Ala Arg Ala Thr Glu
```

RAW SEQUENCE LISTING DATE: 07/12/2006
PATENT APPLICATION: US/10/762,769A TIME: 09:52:58

Input Set : F:\SEQLIST.TXT
Output Set: N:\CRF4\07122006\J762769A.raw

```
58              180                 185                 190
59 Pro Val Ala Met His Ala Tyr Tyr Val Thr Phe Ser Cys Ser Leu Ile
60          195                 200                 205
61 Ala Ala Ala Ile Asn Cys Val Phe Gly Phe Val Leu Ala Trp Val Leu
62      210                 215                 220
63 Val Arg Tyr Asn Phe Ala Gly Lys Lys Ile Leu Asp Ala Ala Val Asp
64 225                 230                 235                 240
65 Leu Pro Phe Ala Leu Pro Thr Ser Val Ala Gly Leu Thr Leu Ala Thr
66                  245                 250                 255
67 Val Tyr Gly Asp Glu Phe Phe Ile Gly Gln Phe Leu Gln Ala Gln Gly
68              260                 265                 270
69 Val Gln Val Val Phe Thr Arg Leu Gly Val Val Ile Ala Met Ile Phe
70          275                 280                 285
71 Val Ser Phe Pro Phe Val Val Arg Thr Met Gln Pro Val Met Gln Glu
72      290                 295                 300
73 Ile Gln Lys Glu Met Glu Glu Ala Ala Trp Ser Leu Gly Ala Ser Gln
74 305                 310                 315                 320
75 Trp Arg Thr Phe Thr Asp Val Val Leu Pro Pro Leu Leu Pro Ala Leu
76                  325                 330                 335
77 Leu Thr Gly Thr Ala Leu Ala Phe Ser Arg Ala Leu Gly Glu Phe Gly
78              340                 345                 350
79 Ser Ile Val Ile Val Ser Ser Asn Phe Ala Phe Lys Asp Leu Ile Ala
80          355                 360                 365
81 Pro Val Leu Ile Phe Gln Cys Leu Glu Gln Tyr Asp Tyr Val Gly Ala
82      370                 375                 380
83 Thr Val Ile Gly Thr Val Leu Leu Leu Ile Ser Leu Val Met Met Leu
84 385                 390                 395                 400
85 Ala Val Asn Gln Leu Gln Lys Leu Ala Arg Lys
86                  405                 410
89 <210> SEQ ID NO: 2
90 <211> LENGTH: 3873
91 <212> TYPE: DNA
92 <213> ORGANISM: Chlamydomonas reinhardtii
94 <400> SEQUENCE: 2
95 gcttagtacc taagcaaaaa taccaaagcc ttatcctgag ttgtcaacaa gaactccagc 60
96 ctgcgacgat gcaaagcctt tcttgagcgg gttgatggac tttgctttgt tatctgtcca 120
97 gtaagccacc agacactacc aagtagagta atccatttgt ataggtacag aatatggagc 180
98 gagtttgcag ccatcagctt gcctcgtcgc gagggaggcc atgcatcgct ggggtgcagc 240
99 ggtcgcccat ccgactaggg acttcaagcg ttgctcatgt gcaggtctct ccggcaggta 300
100 agcaccgcgc tcggcggcgt gtacacatgg ggccgtcagg ccaactgcgt ttgttggcta 360
101 tgcaaccgaa acaggccttg ggagatatca acggcaaaga ctgcaagtcg tggcgtctgc 420
102 agctgcggca gcggctttcg accctcctgg aggtgcgtgg cgtgagggct gcacgggtgc 480
103 gggttggcct ggaaaccaag cctcgccacg actacctgca acagcattgc ccgcatctcc 540
104 agcccctcac cctcgagtgc ctcccgaaga cctctatccc ctgcgcatca ttggttcggg 600
105 ggcgccgcct gcgggccttg ggcgctggct acgctgaccg cacggcacga cttggcacgg 660
106 cctggcgcgg cctgagcggc ccccccctc ctgatggccc cacgctttgc cgcccacgcc 720
107 gctccccgca ggtgtctccg ccgggttctc gcagccgcaa cagcagctgc cacaacagca 780
108 cccacgccaa ccacaggcgg tggcggaggt agctgtcgcc gagtcagtct cggcgcccgc 840
109 ttctgcggcg ccctccaatg atggctcgcc cacggcctcc atggacggcg gccccagctc 900
```

RAW SEQUENCE LISTING DATE: 07/12/2006
PATENT APPLICATION: **US/10/762,769A** TIME: 09:52:58

Input Set : **F:\SEQLIST.TXT**
Output Set: **N:\CRF4\07122006\J762769A.raw**

110 cggcctcagc gccgtgcccg ccgccgccac cgccaccgac ctcttctccg ccgcggcgcg 960
111 cctccgcctg cccaacctct cccccatcat cacctggacc ttcatgctct cctacatggc 1020
112 cttcatgctc atcatgccca tcaccgcgct gctgcaaaaa gcctcgctcg tgccgctcaa 1080
113 cgtcttcatc gcgcgcgcca ccgagccggt ggcgatgcac gcctactacg tcaccttctc 1140
114 ctgctcgctg atcgcggccg ccatcaactg cgtgtttggc ttcgtgctgg cctgggtgct 1200
115 ggtgcgctac aatttcgcgg ggaagaagat cctggacgcg gcggtggacc tgccgttcgc 1260
116 gctgccgacc tcggtggcgg gcctcacgct tgccacggtg tacggcgacg agttcttcat 1320
117 cggccagttc ctgcaggcgc agggcgtgca ggtgcgtgcg tatagcatag tggagtgtgg 1380
118 ttagcagctg ggggtccggc agtagttccc gccctagtga ggtcgaaact ataccagaag 1440
119 aagaggacga acatggggct atccagcaag ctcgtctagg gaaggaggag tttgggagaa 1500
120 cggtggggtg ggagggagag ggagggcgtt ggctgggagg gaagggtaag gcgggaggga 1560
121 gatggtagca cggggcgttg gggacgcaga aggatgacag gcggctgcag ggaagggatg 1620
122 gggaagcgga gctggggaca gtgcgaagag ccgggagaga ggggaagttt gagtcaggaa 1680
123 gaggggctag agaggggcat gcggactcct gctgggattt aggtgcgtgc tcattgagga 1740
124 gcccttggaa tcagcggacg gaaacgtggc cgacggggtc tgccgagcac accaggctag 1800
125 ctagacgcgc ggttgggcaa cgagcagagc tgctgtgcgg ctatggatgg aaggcgatgc 1860
126 agcgagcatg tgcagtgaac attggtttga ggacagggga ctccgaggtt gcataggcgg 1920
127 gccgccactg tctctgccgc tagggtgact agctgcctcg aacctggcgg tggccccata 1980
128 cccgcagttg gaggatgctc cacgcgcttc agcttgccat gtctggggtc tgggtctgga 2040
129 cgcaatcagc gtgtgagggt ccaactctat atggaattat ggataccttc caactaccag 2100
130 cacgtaggct gccggaacgc ggctgaagcg gctggcctgc cccctcatcc tctcgttccc 2160
131 ctgtttttgt cccctgtcca cccaggtggt gttcacgcgg ctgggtgtgg tgatcgccat 2220
132 gatcttcgtg tccttcccct tcgtggtgcg caccatgcag cccgtcatgc aggtgagagc 2280
133 gcccaggagg cggagccatg gcgggttggg gcgggttggg gcgggttggg gcggggcgcg 2340
134 gatggggcgg cttggggagt aatgtggggc ggatggggtg gcagcctggc agggtatggg 2400
135 agcgagagga tagcggggac aggggacagg gaagggaagg gaaggggaag gatgccctat 2460
136 gcgagcaaag ggggtatggg aaccggcggt tggggctggg agcgacggga gcagggaggg 2520
137 agtgcacgga acgggggcaa ggcggacagg gtgagggagg gtgcaggccg gactgggatg 2580
138 ggtcatgtgt cctggtcggg ggtgtagccg tgggaggcgg gcaggcagcg tgtgttctgg 2640
139 cacggtgttt tggcgaaaga taccacggca tggtatgggg ccagttgggc agggaagaac 2700
140 cgttggacac gacttcgttg acagatctag ttcattgcac ccgggtcgca ccaagggtgg 2760
141 cggcgagccc ggcccggcac gtccgagtac cccggagccg taacgccgca acccgccttg 2820
142 ttgcgcccct tccctgctcc cctgctccgc ataccgtgca ccatgccctc tgccgccccc 2880
143 tcaggccctc aggccctcac ctccccctca cctcctccta acgccttccc ctcgccttcc 2940
144 cttcccctcc caacgccacc acgtgcaaca ggaaatccaa aaggagatgg aggaggcggc 3000
145 atggtcgctg ggcgcctcgc agtggcgcac cttcacagac gtggtgctgc cgccgctgct 3060
146 gcccgcgctg ctgaccggca cggcactggc cttctcgcgc gcgcttggcg agttcggatc 3120
147 cattgtcatc gtgtcctcca actttgcctt caaggacctg atcgcgcccg tgctgatctt 3180
148 ccagtgcctg gagcagtacg actacgtggg cgccaccgtg atcggcacag tactgctgtt 3240
149 gatttcgctg gtgatgatgt tggcggtgaa ccagctgcag aagctggcgc gcaagtgagg 3300
150 ggctgaggcg tttgaggaga gtgggcgtct gcggaggcgc ttgtggcgca ggggcaggtg 3360
151 gaggaggttg cagggtgagg caggagtggc aggtggtgga gggtgcaggg cggggtgttg 3420
152 ggatgggatg ggatgggacc gtgggagggg tgggactttg ggtgggtggg agtgggtgct 3480
153 acgtattagg atatgggagg tggtatgcag ttgaaggggg gggtggcaat ctggacgggg 3540
154 actcactgtt tactaggcac gcatgtcgca ggagtggata tcgatgggtg tggggatgtc 3600
155 agcacgcttg gcttgagttg ggccatggga cccgggacta ggcttggttg cgagccgagc 3660
156 cagtcaccag ggagacgtac gagcgcacac agtgattacg gggattgatt aggcggcgaa 3720
157 ttgacgcaaa tccacggggg ctgtggcttg ggggaggcag ggattgagcg aaggacgcac 3780
158 tgcaagctca ggcagtcgca tgcccgtacc ctgcttctgg tccagtgtgg agacaagact 3840

RAW SEQUENCE LISTING DATE: 07/12/2006
PATENT APPLICATION: US/10/762,769A TIME: 09:52:58

Input Set : F:\SEQLIST.TXT
Output Set: N:\CRF4\07122006\J762769A.raw

ggcaatcgtg gtcctttgca attcatggcg cgc 3873

<210> SEQ ID NO: 3
<211> LENGTH: 1984
<212> TYPE: DNA
<213> ORGANISM: Chlamydomonas reinhardtii

<400> SEQUENCE: 3

gcttagtacc taagcaaaaa taccaaagcc ttatcctgag ttgtcaacaa gaactccagc 60
ctgcgacgat gcaaagcctt tcttgagcgg gttgatggac tttgctttgt tatctgtcca 120
gtaagccacc agacactacc aagtagagta atccatttgt ataggtacag aatatggagc 180
gagtttgcag ccatcagctt gcctcgtcgc gagggaggcc atgcatcgct ggggtgcagc 240
ggtcgcccat ccgactaggg acttcaagcg ttgctcatgt gcaggtctct ccggcaggcc 300
ttgggagata tcaacggcaa agactgcaag tcgtggcgtc tgcagctgcg gcagcggctt 360
tcgaccctcc tggaggtgtc tccgccgggt tctcgcagcc gcaacagcag ctgccacaac 420
agcacccacg ccaaccacag gcggtggcgg aggtagctgt cgccgagtca gtctcggcgc 480
ccgcttctgc ggcgccctcc aatgatggct cgcccacggc ctccatggac ggcggcccca 540
gctccggcct cagcgccgtg cccgccgccg ccaccgccac cgacctcttc tccgccgcgg 600
cgcgcctccg cctgcccaac ctctccccca tcatcacctg gaccttcatg ctctcctaca 660
tggccttcat gctcatcatg cccatcaccg cgctgctgca aaaagcctcg ctcgtgccgc 720
tcaacgtctt catcgcgcgc gccaccgagc cggtggcgat gcacgcctac tacgtcacct 780
tctcctgctc gctgatcgcg gccgccatca actgcgtgtt tggcttcgtg ctggcctggg 840
tgctggtgcg ctacaatttc gcggggaaga agatcctgga cgcggcggtg gacctgccgt 900
tcgcgctgcc gacctcggtg gcgggcctca cgcttgccac ggtgtacggc gacgagttct 960
tcatcggcca gttcctgcag gcgcagggcg tgcaggtggt gttcacgcgg ctgggtgtgg 1020
tgatcgccat gatcttcgtg tccttccgct tcgtggtgcg caccatgcag cccgtcatgc 1080
aggaaatcca aaaggagatg gaggaggcgg catggtcgct gggcgcctcg cagtggcgca 1140
ccttcacaga cgtggtgctg ccgccgctgc tgcccgcgct gctgaccggc acggcactgg 1200
ccttctcgcg cgcgcttggc gagttcggat ccattgtcat cgtgtcctcc aactttgcct 1260
tcaaggacct gatcgcgccc gtgctgatct tccagtgcct ggagcagtac gactacgtgg 1320
gcgccaccgt gatcggcaca gtactgctgt tgatttcgct ggtgatgatg ttggcggtga 1380
accagctgca gaagctggcg cgcaagtgag gggctgaggc gtttgaggag agtgggcgtc 1440
tgcggaggcg cttgtggcgc aggggcaggt ggaggaggtt gcagggtgag gcaggagtgg 1500
caggtggtgg agggtgcagg gcggggtgtt gggatgggat gggatgggac cgtgggaggg 1560
gtgggacttt gggtgggtgg gagtgggtgc tacgtattag gatatgggag gtggtatgca 1620
gttgaagggg ggggtggcaa tctggacggg gactcactgt ttactaggca cgcatgtcgc 1680
aggagtggat atcgatgggt gtggggatgt cagcacgctt ggcttgagtt gggccatggg 1740
acccgggact aggcttggtt gcgagccgag ccagtcacca gggagacgta cgagcgcaca 1800
cagtgattac ggggattgat taggcggcga attgacgcaa atccacgggg gctgtggctt 1860
gggggaggca gggattgagc gaaggacgca ctgcaagctc aggcagtcgc atgcccgtac 1920
cctgcttctg gtccagtgtg gagacaagac tggcaatcgt ggtcctttgc aattcatggc 1980
gcgc 1984

<210> SEQ ID NO: 4
<211> LENGTH: 1863
<212> TYPE: DNA
<213> ORGANISM: Chlamydomonas reinhardtii

<400> SEQUENCE: 4

cattcaattt gcagcgttcc taaaatggca agcacaacgc tgctccagcc cgcgcttggt 60
ctgccctcgc gggtagggcc tcgctcccct ctgtcgcttc ccaaaattcc tcgcgtgtgc 120
acgcacacta gtgctccctc tacctcaaag tactgcgact catcatcagt tatagagagc 180
acgctagggc ggcaaacatc ggttgccggg agaccatggc ttgcaccccg gcctgcgcct 240

RAW SEQUENCE LISTING DATE: 07/12/2006
PATENT APPLICATION: US/10/762,769A TIME: 09:52:58

Input Set : F:\SEQLIST.TXT
Output Set: N:\CRF4\07122006\J762769A.raw

212 caacaaagcc gaggcgacct actggtctcc aaatcggggg cagcaggagg catgggcgcc 300
213 catggagggg gcttagggga accggtcgat aattggatca agaagctact cgttggtgtc 360
214 gcggcggcgt acatcggctt ggtcgtgctg gtgcccttcc tgaatgtctt cgtccaggcg 420
215 ttcgccaagg gcatcattcc cttcctggag cactgcgcgg acccggactt tctgcacgca 480
216 ctcaagatga cgctgatgct ggcgttcgtg acggtgccgc tcaacacggt gtttggcacg 540
217 gtggccgcga tcaacctcac gcgcaacgag ttccccggca aggtgttcct gatgtcgctg 600
218 ctggacctgc ccttctccat ctcgcccgtg gtgactggcc tgatgctcac gctgctgtac 660
219 ggccgcaccg gctggttcgc ggcgctgctg cgggagaccg gcatcaacgt ggtgttcgca 720
220 ttcacgggca tggccctggc caccatgttt gtgacgctgc cgttcgtggt gcgcgagctg 780
221 atccccatcc tggagaacat ggacctgtcg caggaggagg cggcgagaac gctgggggcc 840
222 aacgactggc aggtgttctg gaacgtgacg ctgcccaaca tccgctgggg cctgctgtac 900
223 ggcgtgatcc tgtgcaacgc ccgagccatg ggcgagttcg gagccgtgtc cgtcatctcg 960
224 ggcaacatca tcggccgcac gcagacgctg acgctgttcg tcgagtccgc ctacaaggag 1020
225 tacaacacgg aggcggcgtt cgcggcggct gtgctgctga gcgcgctggc gctgggcacc 1080
226 ctgtggatca aggacaaggt ggaggaggcg gcggcggcgg agagccgcaa gtagagagga 1140
227 gcaggcggcg tcggcagcgg cggcagtggc agcggcagcg gcggagagcg gcagctggag 1200
228 aggagcaggc ggtggcggcg gagcggcgga aatagagagg tgcagcaagg aggcaggcgc 1260
229 cgacgcgagg ggagggcgtg gtggtgggct tgcgtgggtg cttggtccgt ggccagggtg 1320
230 cctggcctgg gtagttggtg tgtgggtgaa gctgattcct gtttgggtga ggcggccgag 1380
231 ttcctgaagg aagcaaggaa ggacagtgcc gcagtgacca gcgggtaatg gtaagggagc 1440
232 tgacacgtgt ggcgttctgt tgctggtcgc cgcatgctta acgcagcggg agcagcttct 1500
233 ctgtctgatg tctaacgggg gcgttgtatg ctgataatag acggagggcg aagggagcag 1560
234 gcgcggttca gatggggtaa aagctgttgg aaatcaacac gtgcagcggg tgggttgcat 1620
235 ttgtgatcac tggacgttct gagtggtccg tgcgcctata gcgcgtgctg tgcatatata 1680
236 cgcgcgccgg cgcataaaac atgactgcat gtgtcggtgt tgacggtaca gttatgccgt 1740
237 gccccgtttt acaagcggga tagaggcaca ctccacgtag tatgcattga gcccagtaga 1800
238 ctctggtcag aaggccggta aatttacatg tgtcgtggtg aaccctgtaa gtcatggccc 1860
239 aag 1863
241 <210> SEQ ID NO: 5
242 <211> LENGTH: 2253
243 <212> TYPE: DNA
244 <213> ORGANISM: Chlamydomonas reinhardtii
246 <400> SEQUENCE: 5
247 gtacttcaat tgtcagaatg gcgtcgctgc tcgctcaaac aacatcgcgc cttggcgctc 60
248 gcccagctgc gcaagctggc cctgtcgccc aaatggcacc gatggcaagc cgagtgcagc 120
249 cggcgatgcc tagcgcgctg ctcccactgc acgccagagc gacaacaact tcagtcgctt 180
250 gccgggcagc cagcatcgac aaacctgtcg tttacactcc tcgagattcg tcgcaacagt 240
251 cctccaatgg ggcaggagaa gtgtccatgt ccatatcatc catggacgag gttggaccct 300
252 cttatgaggg aatcattaca gacgcgccta cacgaccaac ggggctttat gtgcgggtgc 360
253 gcaacatggt gaagcacttc agcaccgcca aaggcctgtt cagggcggtg gacggcgtgg 420
254 acgtggacat cgagcccagc tccatcgtgg cgctgctggg gcccagcggc agcggcaaga 480
255 ccacattgct gcgcctcatt gcaggcctgg agcagcccac gggcggcaac atctactttg 540
256 acgacacgga cgcgaccaac ctgtccgtcc aggaccgcca gatcggcttc gtgttccaga 600
257 gctatgcgct gttcaaccac aagacagttg cggagaacat caagtttgga ctggaggtgc 660
258 gcaagctcaa catcgaccac gacaagcgcg tggcggagct gctggcgctg gtgcagctca 720
259 ccggcctggg cgaccgctac ccgcgccaac tgtcgggcgg ccagcggcag cgtgtggcgc 780
260 tggcgcgcgc cctggcctcc aacccgcggc tgctgctgct ggacgagccc tttggcgcgc 840
261 tggacgcggt ggtgcgcaag cagctgcgca cggggctgcg cgagatcgtg cgcagcgtgg 900
262 gcgtgaccac catcattgtg acgcacgacc aggaggaggc gttcgacctg gcggacaagg 960

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/12/2006
PATENT APPLICATION: **US/10/762,769A** TIME: 09:52:59

Input Set : **F:\SEQLIST.TXT**
Output Set: **N:\CRF4\07122006\J762769A.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; Xaa Pos. 438

VERIFICATION SUMMARY DATE: 07/12/2006
PATENT APPLICATION: US/10/762,769A TIME: 09:52:59

Input Set : F:\SEQLIST.TXT
Output Set: N:\CRF4\07122006\J762769A.raw

L:445 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:432